DOCUMENT RESUME

ED 073 845 PS 006 381

AUTHOR Regan, Ellen M.; And Others
TITLE Evaluation and Implementation of the Conceptual Skills Program.
PUB DATE Feb 73
NOTE 45p.; Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, February 25 - March 1, 1973

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Achievement Tests; *Concept Formation; *Curriculum Evaluation; *Kindergarten; Parent Attitudes; *Skill Development; Student Attitudes; Teacher Evaluation; Technical Reports; Thought Processes

ABSTRACT

The evaluation of the Conceptual Skills Program, a curriculum program for the kindergarten, and teacher implementation of the program are reported. The evaluation was comprised of three components. The first, an estimate of student achievement, relied mostly on the comparison of the performance of program objectives of kindergarten classes in 12 schools in which the program was initiated in the fall of 1971 with the performance of program objectives of the children who had immediately preceded them in the same schools and who had not received the program. Additional data were collected from the classes of teachers in the area who were teaching the program for the second or third year. Results indicated significantly greater conceptual skills achievement by students exposed to the program. The second component, a student attitude measure, estimated the relative performance of children for Conceptual Skills and seven other in-and-out-of-school activities. Data were collected from all children receiving the program, which necessitated devising a group-administered attitude test for children unable to read or write. Students ranked the Conceptual Skills Program lowest. Opinion data from parents, the third component, were also collected, using an interview-questionnaire technique for a random sample of parents. Generally, parents were well informed about the program and approved of it. Ten teachers were observed and evaluated in their use of the program. (Author/KM)

FILMED FROM BEST AVAILABLE COPY

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

NPD
PS

ED 073845

EVALUATION AND IMPLEMENTATION OF THE CONCEPTUAL SKILLS PROGRAM

Ellen M. Regan
THE ONTARIO INSTITUTE FOR STUDIES IN EDUCATION

Kenneth A. Leithwood
TRENT VALLEY CENTRE - O.I.S.E.

Florence Maynes
TRENT VALLEY CENTRE - O.I.S.E.

PS 006881

Paper presented at AERA meeting in New Orleans, 1973.

The development, field testing, evaluation and implementation of a curriculum program provide considerable and varied data to those involved in this kind of change process. This report which discusses the evaluation of a curriculum program and teacher implementation of the program sheds some light on the process of effecting school change.

The curriculum program considered is the Conceptual Skills Program (Bereiter-Regan) developed for use in the kindergarten. The purpose of this program is to develop skill in the use of simple concepts. It is primarily a thinking program, not a language program, as it is concerned with teaching children to communicate ideas accurately, see relationships among concepts and use concepts effectively in thinking. The learning tasks and instructional materials developed are intended for use in daily lessons which represent only a part of the total kindergarten program.

Program Evaluation

The 1971-72 evaluation of the program conducted by the Trent Valley Centre was comprised of three components. The first, an estimate of student achievement, relied for the most part on the comparison of the performance on the objectives of the program of kindergarten classes in 12 schools where the program was being initiated in the fall of 1971, with the performance on the objectives of the program of the children who had immediately proceeded them in the same schools and who had not received the program. Additional data were collected from the classes of teachers in the area who were teaching the program for the second or third year. Results indicate significantly greater ($p < .001$) conceptual skills achievement on the part of students exposed to the program ($M = 21.42$ $n = 487$) as compared with comparison students ($M = 19.58$ $n = 498$).

The second component, a student attitude measure, estimated the relative performance of children for Conceptual Skills and seven other in-and-out of school activities commonly engaged in by children of this age. These data were collected from all children in the area who received the program and necessitated the devising of a group-administered attitude test which could be used with young children not yet able to read or write. Although the scale used in the evaluation prevented absolute estimates of perference, students ranked the Conceptual Skills Program lowest in relation to activities, all of which may well have been considered enjoyable by them.

Opinion data from parents, the third component, were also collected using an interview-questionnaire technique in which randomly selected parents were visited and a questionnaire completed in their presence by an interviewer. Generally, parents were well informed about the program and approved of it.

The following discussion considers each of the evaluation components in more detail. The measures and procedures used in collecting the data, as well as methods of data analysis, are described.

Student Achievement

To provide comparative data, testing on the objectives of the program was conducted in the fall of 1971 with children who had completed kindergarten and were commencing Grade 1 in schools where the program was just being introduced in kindergarten. The same tests were then re-administered in June, 1972, to the children in the same schools who were then just completing their kindergarten year. This design, the post-test, post-test time series design, permits the control of teacher, school and community variables which affect the interpretability of achievement scores and is discussed in detail in Leithwood and Russell (1972).

Each item in the set of 143 test items designed by the program developers for group testing of the objectives of the program was randomly assigned to one of six sub-tests, and one of these sub-tests was then randomly assigned to each of the control (Grade 1, fall 1971) and treatment (kindergarten, June 1972) classes. The test was administered by the classroom teacher and standarization was achieved through the use of taped instructions to the students and a set of mimeographed directions for the teachers.

In all but one school the same teacher taught both control and treatment children in kindergarten.

Results

The mean test score of the control children (out of a possible 24 and, with one sub-test, 23 mastered objectives) was 19.58 and of the treatment children was 21.42. Analysis of variance showed this difference was highly significant. Children who received the program had, after approximately four months less maturation, development and schooling, outperformed children from the same schools in the same areas with the same kindergarten teachers on the objectives of the program. The difference was consistent for all test forms, and appeared unaffected by some differences in the socioeconomic status of the control and treatment children who received a given test form.

Table 1 summarizes the proportions of correct responses to clusters of items testing the objectives of the program. On most objectives, and on all levels individually and combined, the treatment children performed better than the control children. Inspection of this table also shows particular objectives where the previous instruction was apparently equally

TABLE 1

Proportion Of Correct Responses To Item Clusters Measuring
Program Objectives In Control And Treatment Data

LEVEL	OBJECTIVES (Showing Item Numbers)	CONTROL			TREATMENT		
		#C[1]	N[2]	P[3]	#C[1]	N[2]	P[3]
1	Objects (1 to 4)	262	265	.989	332	347	.958
1	Size (5 to 8)	315	319	.987	321	322	.997
1	Negative (9, 10)	167	188	.888	149	153	.961
1	Shape (11 to 14)	336	396	.848	318	320	.994
1	Location (15 to 16)	208	224	.929	249	253	.984
1	Use (18, 19)	138	141	.979	161	161	1.00
1	Action (20, 21)	248	250	.992	150	150	1.00
1	Parts (22, 23)	146	161	.907	135	164	.823
1	Color (24 to 26)	265	266	.996	235	236	.996
	TOTAL 26 ITEMS	2085	2210	.943	2050	2106	.973
2	Number (1 to 8)	690	696	.991	640	646	.991
2	Location (9 to 17)	571	781	.731	654	726	.901
2	Inst. Following (18 to 22)	246	429	.573	339	396	.856
	TOTAL 22 ITEMS	1497	1896	.790	1633	1768	.924
3	Location (1 to 14)	782	877	.892	837	909	.921
3	Objects (11 to 13)	43	170	.253	209	267	.783
3	Sameness (15 to 20)	491	527	.937	452	478	.946
3	Difference (21 to 26)	378	401	.943	453	488	.928
3	2-Part Inst. Follow (27 to 29)	116	254	.457	175	248	.706
	TOTAL 29 ITEMS	1813	2229	.813	2126	2390	.890
4	Size (1 to 8)	561	577	.972	653	668	.978
4	Comparative (9 to 16)	597	621	.961	632	658	.960
4	Location (17, 18)	69	114	.605	137	168	.815
4	Sameness (19 to 21)	202	214	.941	257	257	.922
4	Difference (22 to 24)	262	286	.916	229	239	.958
4	Incongruity (25 a,b; 26 a,b)	301	352	.855	296	330	.897
	TOTAL 28 ITEMS	1992	2164	.921	2184	2320	.941
5	Definition (1 to 6)	520	616	.844	442	464	.953
5	Consequences (7 to 9)	226	281	.804	219	241	.909
5	Sameness & Difference (10 to 23)	674	1111	.607	839	1117	.751
	TOTAL 23 ITEMS	1420	2008	.707	1500	1822	.823
6	Relevance (1 to 3)	285	298	.956	222	227	.978
6	True-Not True (4 to 6)	175	197	.888	250	260	.962
6	Class Inclusion (7 to 11)	203	449	.452	256	399	.642
6	Sameness (12 to 15)	276	351	.786	235	313	.751
	TOTAL 15 ITEMS	939	1295	.725	963	1199	.803
	GRAND TOTAL ALL LEVELS 143 ITEMS	9746	11802	.826	10456	11605	.901

1 #C - number of correct responses

2 N - number of times the items in a cluster of items were tried.

3 P - proportion

as effective as the program instruction (for example, items 5 to 8 in level one, testing acquisition of the concept of size where, out of 319 tries at the three items 315 correct answers were recorded by the controls, and 321 correct answers out of 322 tries were recorded by the treatment children) and also particular objectives where the previous instruction was much less effective (for example, item 27 to 29 in level 3, testing the objectives of two-part instruction following, where 116 correct responses out of 254 tries were recorded by the controls, a proportion of .457, and 175 out of 248 correct responses were recorded by the treatment children, a proportion of .706).

The results of testing a number of children whose teachers were in their second or third year of teaching the program were also analyzed. Although they also outperformed the control group, they did less well than the treatment group. This result was thought to indicate that longer experience with teaching the program does not necessarily improve, and is not necessary to effective instruction on the program objectives. However, the data analysis also showed that the socioeconomic status of the children in the classes of second year teachers was lower than the treatment group and also lower, though not significantly so, than the control group. Ability differences could conceivably accompany this S.E.S. difference and have affected the performance of the second year teachers' students.

PS 006381

Student Attitude

The test devised to obtain data on this component of the evaluation required only that the children examine, two at a time, projected photographs of children of their own age engaged in in-and out-of-school activities typical

for children of their age. With each pair they were asked to choose which they would rather do, and to mark the square on their answer book page on the same side as the pictured activity in which they would prefer to engage. The test design permitted a comparison of each of the activities once with every other one (there were 8 activities and a total of 28 choices) and the resultant data permitted the calculation of scale values for each activity which estimated not only the ranking the children assigned to the activities, but the distances between the values. Paired comparison scaling data analysis methods are outlined in Edwards (1957).

The photographed activities are listed below:

1. Conceptual Skills
2. Singing with the class
3. Show and Tell
4. Painting
5. Play with toys and blocks
6. Being read a story
7. Outdoor play
8. Watching TV

Each photograph showed several children with a young female adult in brightly coloured clothing. In six of the photographs, including Conceptual Skills, the setting was a classroom. For the last two photographs, a home setting was chosen. Watching TV was photographed in the living room and Outdoor Play, a winter scene involving skating, sleds and a snowmobile, in the yard of a home. The photographs were presented by means of two slide projectors and for the sake of standardization the instructions to the children were placed on tape.

All classes where the program was in operation were tested. All data were checked against criteria of consistence and eliminated if it appeared that the child had been making choices more or less at random. (Perfect consistency is demonstrated by evidence that the favourite activity is chosen in all of the pairs in which it appears, the second rank in all but one and so on to the least preferred which is never chosen.) The rejected data came from children whose achievement test scores were somewhat, but not greatly, lower than that of the children whose data were accepted.

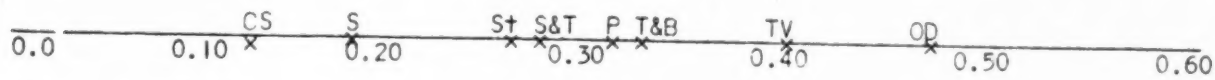
The accepted data were grouped according to three dimensions which, it was considered, could affect the attitude of the children. These were the experience with the program of the teachers (first year vs. two or three years), the experience of the children with a school environment (children who had attended nursery school or, following a year of kindergarten were taking the program in a K-1 class, vs. those having their first experience with school) and the judged competence with the program of the teachers involved.

Results

Figure 1 shows the ranking produced by the total group and by each of the breakdowns noted above on a scale of normal deviates. It is apparent that the various breakdowns produce scalings which show little variation from each other, or from the scale values obtained using the total set of data. Outdoor Play is most preferred, followed by TV and the within-school activities ranked in a continuum which appears generally to be characterized by the amount of structure and direction inherent in the activities themselves. Conceptual Skills then is the least preferred activity among the eight presented to the children.

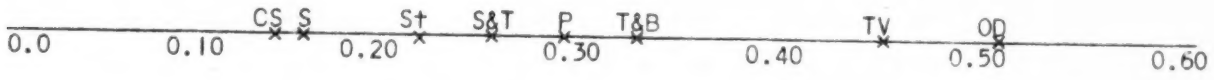
LEGEND: Conceptual Skills - CS
Singing - S
Story - St
Painting - P
Show and Tell - S&T
Toys and Blocks - T&B
Television - TV
Outdoor Play - OD

All Children (N=398)



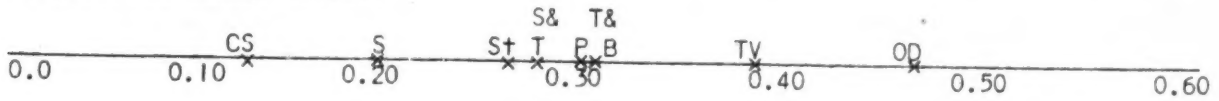
mean normal deviate z + .300

Previous school (N+58)



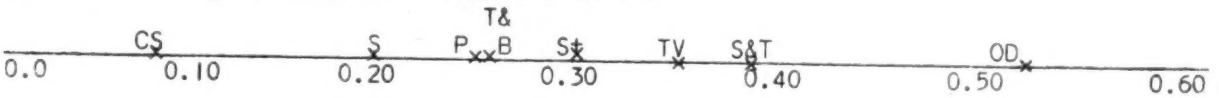
mean normal deviate z + .300

No previous school (N=340)



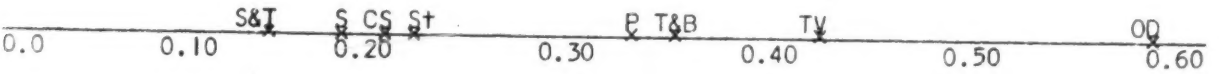
mean normal deviate z+.300

Students of highly rated C.S. teachers (N=77)



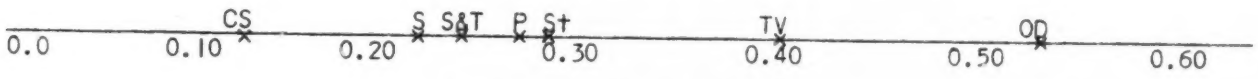
mean normal deviate z+.300

Students of below average C.S. teachers (N=95)



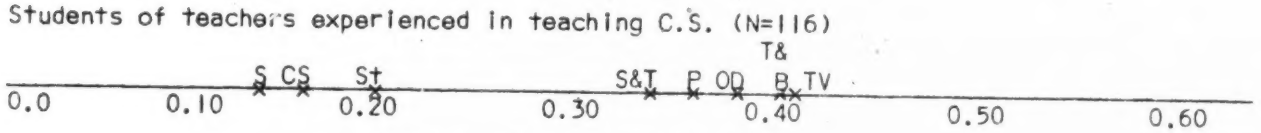
mean normal deviate z+.300

Students of teachers new to C.S. (N=282)



mean normal deviate z+.300

Students of teachers experienced in teaching C.S. (N=116)



mean normal deviate z+.300

FIGURE 1 - Scaled ordering of 8 selected kindergarten-age activities (including Conceptual Skills) for divisions of the sample and for the sample overall.

In interpreting this result it should be kept in mind that this scale does not in itself indicate the rejection or approval of any of the activities, but only their ranking relative to each other. It may also be pointed out that the method produces a scale value for any activity which is a function of the proportion of children who have chosen it over all the others. Hence if most of the children had chosen one of them, and most had never chosen another, the scale itself would be several units long. In this case, all the activities are clustered within a space of 0.60 unit normal deviates, i.e., less than one of the units. During the data analysis, it was evident that with most choices (i.e., 23 out of the 28) the less popular activity was still chose by over 40% of the children and this is reflected in the short distance between the most preferred activity, Outdoor Play, and the least preferred, Conceptual Skills. In other words, while individual children have expressed reasonably consistent preferences, there is no indication in the data of a very strong order of perference which is common to most of them.

A test-retest reliability estimate was also made by repeating the test one week later in two schools. The correlation between number of choices of Conceptual Skills on the two occasions was 0.55.

Parent Attitude

The parents of six randomly selected children in each of the two "old" (second or third year of the program) and two "new" (first year of the program) schools were interviewed. Generally parents were quite well informed about the program and approved of it. The majority approved of the idea of formal instruction being introduced in kindergarten. Reasons for their approval varied with the parents' preconceived ideas of what constituted "teaching" i.e., formal instruction.

Thus the "progressive" parent willing to explore new ways and the strict disciplinarian, who differentiates between the "learning" and "play" activities in school, were satisfied with the content and structure of Conceptual Skills Program. About two-thirds of the parents were unable to determine accurately whether the Conceptual Skills Program had affected their children's learning or achievement, or their attitude towards the school. Although most of them discussed the worksheets with their children, many could not quote verbatim comments made by the child about his special work. Yet one half felt that the children approved of (liked) the Conceptual Skills Program. Only one registered "somewhat disapproving" attitude by their child.

Most parents saw no objection to the use of programs which have been developed by experts outside this area. Their main concern was the quality and relevance of the program.

Summary

The findings related to student achievement are generally consistent with finding of earlier evaluations (1968-69, Bereiter-Scardamalia; 1970-71, Russell and Leithwood). However this evaluation represented the first attempt to obtain specific evidence of student and parent attitude.

Teacher attitudes, sampled by questionnaire throughout the period of field trials, were generally positive. However sustained contact with teacher users during the field trial period revealed certain factors critical to successful program implementation.

Preparing Teachers for Program Implementation

Joyce (1969) observes that the acceptance and life span of any innovation is largely dependent on how competent teachers feel in using or implementing the

innovation. Teachers' feelings of competence are undoubtedly influenced by the extent to which they feel prepared to cope with the change inherent in implementing new programs and practices. As Eisner (1970) states, "New curricula often demand of teachers a new way of looking at the task of instruction, and often require that they deal with concepts and procedures that are as new for them as they are for students." (p. 8) This observation is applicable to demands placed on teacher users of Conceptual Skills. Although the simple concepts considered in the program are not "new", the way in which the program seeks to develop these concepts, and skills in using these concepts, does differ from more traditional approaches to these objectives. The recommended instructional techniques are particularly "new" to many teachers for whom direct and systematic instruction represents a new way of working with kindergarten children.

Recognizing what the program expected of teachers, developers attempted to identify the problems experienced by teachers implementing the program and, subsequently, to design an approach to inservice training which could deal with these problems. Thus during four years of field trials, developers maintained contact with all Conceptual Skills teachers through inservice training sessions and regularly scheduled visits to classrooms. There was also an attempt to identify, during the fourth year of field trials, the characteristics of teaching styles and kindergarten programs which seemed to distinguish between more and less successful implementation of Conceptual Skills.

Developing an Approach to Inservice Training

For 4 years, all Conceptual Skills teachers participated in a 2-3 day training period prior to implementing the program. A major objective of this training was developing teacher understanding of the program's purpose and

objectives and how tasks, materials, and instructional techniques were designed to accomplish program objectives. Nevertheless the problems experienced by teachers during the early field trials related in one way or another to poor or limited understanding of the purpose and objectives of the program. Observations in classrooms revealed, for example, major difficulty in planning appropriate daily lessons. Teachers admitted that they were not sure how to select and sequence tasks in a manner that provided for the sequential development of skills.

Lack of understanding resulted also in establishing inappropriate criteria for evaluating student performance and response. Frequently the criteria established were extraneous to the objectives of the program (e.g., children were evaluated on how well they colored the picture of the big dog rather than on the ability to discriminate between the big dog and the small dog).

Failure to implement recommended teaching procedures was a further indication of limited understanding of what the program was designed to accomplish. If, for example, a child could identify a picture as "having something wrong with it", (e.g., tree with square apples) many teachers were satisfied. They did not ask children to explain what was wrong or why it was wrong.... explanations which represent the real purpose of the tasks.

Ironically, what seemed to be interfering with teacher understanding was the fact that many teachers thought they did understand program purpose and objectives. Some suggested that Conceptual Skills was "just another way of doing what I already do in my program." This conclusion seemed to result from teachers responding to what they perceived as the "familiar" without considering what was "new" or "different." Admittedly the simple concepts considered were not "new" to teachers and many tasks appeared similar to traditional kindergarten

activities. In addition, teachers found themselves in agreement with the broad program goals of teaching children to "think" and to "communicate." These broad goals were, of course, compatible with those which teachers had established for their classroom program.

However it was soon evident that in considering program purposes and objectives, many teachers could not discuss or consider these items in other than fairly global terms. Teachers talked about the importance of "children learning to express themselves" and of "teachers providing for individual differences." Yet it was difficult for teachers to talk in specific terms about specific skills involved in "expressing oneself" or specific provisions for particular "individual differences." However teachers did suggest that the Conceptual Skills program manual would be more helpful if it provided more guidance and direction for the teacher. Teachers stated also that developers visits to classrooms and the subsequent assistance provided was the "best" feature of the program.

Thus what teachers "said" and "did" relative to implementing Conceptual Skills supplied developers with clues as to what was interferring with teacher understanding of program purpose and objectives. Further the sustained contact with all teacher users indicated what teachers considered important in developing their competency with the program. All of this feedback served as input to revising the program and to developing an inservice training package.

The Program Manual

It is impossible to consider what was learned about what teachers wanted and needed without reference to the program manual. The first two editions of this manual did not provide sufficient guidance or structure for teachers. Although these manuals organized tasks according to levels

of difficulty (simple to complex thinking operations), day to day program planning was the teacher's responsibility. However the final edition of the manual was structured to the point of outlining a sequence of daily lessons within each program level. As might be expected, teachers using the program for the first time tended to follow the manual closely. During the second year, these teachers did not rely as extensivily on the manual and many began developing their "own" tasks and materials for use in Conceptual Skills lessons. In the words of one teacher, "I really learned about the program from following the manual during the first year." This sentiment was expressed by many teachers and suggests that program manuals can be designed to provide an effective and sustained training input for teachers implementing a new program.

What was observed, and supported in teacher statements, suggests that it may be realistic to think of the first encounter with a new program as a learning experience for the teacher. As the teacher is learning and finding his way, he may not be accommodating to group or individual need to the extent desired. However if the new program is basically sound and if teachers are provided with guidance, it is likely children will benefit. Children in Conceptual Skills classrooms showed, consistently, gains in concept learning and skill mastery.

The Training Package

The attempt to develop ways and means of promoting better teacher understanding of the program's purpose and objectives resulted in the development of a "training package." The package consists of two films and a manual which suggests the format and procedures for three consecutive

training sessions. The manual includes, as well, materials to be reproduced for teacher participants. The films and all other materials focus on analyzing learning tasks, materials and instructional techniques in relation to the purpose and objectives of the program.

Each film has a particular objective but both show teachers working with children in the context of the Conceptual Skills Program. The first film, "Explaining the Conceptual Skills Program," demonstrates the progression of learning tasks from learning basic meanings to applying concepts in increasingly more complex tasks. Attention is given as well to how the materials being used support the purpose of the task. The film is developed in a way which permits "stopping" the film at given points for teacher comment and discussion. In fact the training manual suggests appropriate "stopping" points and provides questions to be considered in analyzing a particular segment (e.g., What are the communication skills that the program seeks to develop?) The training manual also "spells out" procedures for using the program manual to elaborate on information presented in the film. In studying the film and program manual, particular attention is given to involving teachers in a consideration of why a given task is more or less complex than another (e.g., What does a child need to know to handle this task? If children had difficulty with this task, you could drop back to an easier one; which one would you choose?)

In addition to suggestions for considering specific tasks and materials, the training activities built around the first film include suggestions for involving teachers in discussing how Conceptual Skills contrasts with the approach to concept development which the teacher has been using. How is it the same? How is it different?

The second film, "Teaching the Conceptual Skills Program", focuses on developing an understanding of the purpose of recommended teaching techniques. This film, as the first, was developed to permit study of what is observed. Again, the training manual provides suggestions for using the film for given purposes (e.g., discussion questions such as "What did Billy's response indicate to the teacher?"; "What steps did the teacher take in remediating this error?" "How did the teacher probe Lucy's understanding in this task?")

As a supplement to this film, the training manual includes a set of simulation materials. These materials present teachers with problems and situations to consider and discuss. One set of "problems", which describe "individual differences" in understanding and behavior, require teachers to answer the question, "What could (or should) the teacher do in this situation?" A second set of "problems" require teachers to suggest how they would go about remediating specific errors in children's responses (e.g., child who is confusing "left" and "right").

Because the films and other training materials were developed during the third year of field trials, it was possible to "test" their effectiveness during the fourth year of program trials. Teacher response in training sessions and follow up observations in the classrooms of "new" teachers indicated that the materials and procedures encompassed in the training package were effective in "getting at" some of the previously experienced problems of many "new" teacher users of the program.

Assessing Teacher Implimentation

The improvement in teacher training and the sustained contact between the developers and users of the program resulted in increasingly more effective

program implementation. Nevertheless there were indications that the program was "going better" in some classrooms than in others. In some classrooms, lessons "moved along" with the obvious interest and involvement of all children. In other classrooms, a more "plodding" pace was observed with response opportunities provided for fewer children.

There were also indications that in some classrooms, more than in others, Conceptual Skills lessons represented a better "fit" with the total classroom program. For one thing, the ease with which some children approached Conceptual Skills lessons suggested that these lessons did not set expectations or require behavior that differed significantly from what was expected or required in other aspects of their daily program. Similarly some teachers, more than others, reported incidents of how what was considered in Conceptual Skills lessons "transferred" (teachers' term) to other activities carried forward in the classroom. These observations and others suggested that the implementation of Conceptual Skills was influenced by the kind of expectations, objectives and program priorities which teachers established for their total program. Although it is widely recognized and accepted that this is the case in implementing any innovative program or practice, there was interest in determining what factors or conditions discriminated between more and less successful implementation of the Conceptual Skills Program. To this end, a small classroom observation study was carried forward.

The programs of ten teachers were identified for study. The manpower available dictated the number of classroom programs that could be studied in any detail. In selecting the ten classrooms attention was given to including examples of what represented in developers' judgments "good", "adequate" and "poor" implementation of the Conceptual Skills Program. The teachers in this sample

included those "experienced" with Conceptual Skills and those using it for the first time. Five half day observations were scheduled in the classroom of each teacher. Each observation included observing and assessing teacher performance during the Conceptual Skills lesson and observing and describing selected aspects of the remaining classroom program. The same two members of the development team (observer 1 and observer 2) were responsible for all observations in all classrooms. Both observers were involved in assessing teacher performance during the Conceptual Skills lesson; observer 2 was responsible for observing and describing the remaining daily program in all classrooms.

Figure 2 presents the Teacher Evaluation rating form used in assessing teacher performance during Conceptual Skills lessons. Appendix A contains the guide and criteria used in determining the ratings for each of the six items. Both observers observed the same teacher simultaneously but each observer completed his ranking independently of the other observer. Results supported the hypothesized range of teacher implementation (poor-good) considered to be present in the sample. As revealed in the data, there was no necessary correlation in this sample between "experience" and the degree of success in implementing the program. The data indicated, also, that a teacher tended to achieve a higher or lower rating on all items suggesting that the cluster of behaviors identified on the rating form was more or less easily implemented by a given teacher. These results suggest that the teaching behaviors identified on the rating form were more compatible with the established teaching "style" or behavior of some teachers than others.

Figure 3 lists the items considered on the Observation Record used in describing the remaining classroom program. Appendix B contains a sample Observation Record which presents the information recorded in an actual

FIGURE 2

Teacher Evaluation - Conceptual Skills Program

The teacher involves the entire group by -

-- setting the expectation of everyone participating, listening and thinking
-- giving the group time to think before naming whoever is to respond
-- varying group and individual responses
-- requiring everyone to do his own work.

The teacher supports her children as needed by -

-- modelling new work
-- verifying work
-- checking the children's work
-- stepping in to clear up misunderstandings thereby not allowing a child or the group to practice errors.

The teacher gives her children responsibility by -

-- clearly setting expectations and following through positively
-- ceasing to provide a model as soon as the children can work correctly without it
-- giving the children many opportunities to talk
-- asking the children to evaluate their work
-- asking the children to justify their choices (whenever appropriate)
-- asking the children to think of other possibilities (whenever appropriate)
-- requiring the children to handle materials quickly and in response to specific instructions.

The teacher enables the children to work at a success level by -

-- working at a level of difficulty that is challenging to most of the children
-- varying expectations for individual children according to their abilities
-- probing the children's understanding by careful questioning to get at all the elements of the task
-- trying to find a reason for copying and trying to eliminate the cause.

The teacher remediates effectively by -

-- spotting errors that are made
-- determining the cause of errors and leading the children to understanding
-- modelling only when absolutely necessary
-- remediating quickly so as not to "lose" the group
-- providing practice for specific difficulties
-- keeping the remediation positive (emphasizing the idea of working hard to get correct as opposed to the idea that an error has been made).

The teacher understands the intent of each task and clearly works to achieve its purpose.

FIGURE 3

Observation Record

1. Schedule of activities observed.

2. Proportion of total group, small group, individual activity.

3. Nature of tasks, materials etc. for individual and/or small group activity (e.g., what is available for children to choose and/or what does teacher outline for individual or small group activity).

4. Selection and performance of individual and/or small group tasks (e.g., is choice child selected or teacher selected, proportion of individual to group activity).

5. Teacher role during individual/small group activity (e.g., what does the teacher do, how does she interact with children).

6. Nature of tasks for total group activity (exclusive of Conceptual Skills).

7. Selection and performance of tasks in total group activities (e.g., who decides what to "do," can child decide not to participate).

8. Teacher role in total group activities (e.g., what does the teacher do, how does she interact with children).

observation. When the descriptions collected on the Observation Record were compared with the ratings obtained on the Teacher Evaluation forms, the comparison revealed that all items on the Observation Record discriminated to some extent between programs of the 5 teachers achieving "higher" rating on the Teacher Evaluation form and the 5 teachers achieving the "lower" ratings. Comparison of the programs of "higher" and "lower" rated teachers indicated differences in 1) the selection and ordering of activities for the daily program 2) the program priorities which seemed to be established 3) the expectations which seemed to be held for children.

The following discusses the limitations of the study, describes the analysis of the data, and reports the findings in more detail.

Limitations of the Study

The instruments used in collecting the data suffer from certain obvious limitations. The four points rating scale used in the Teacher Evaluation form does not properly reveal either the extremes or the range of differences between a rating of 1 and 4. Nevertheless since it does permit discrimination between satisfactory (adequate-good) and unsatisfactory (weak-poor) teacher performance during Conceptual Skills lessons, it was considered adequate for the purposes of the study.

The Observation Record also suffers from limitations but again it was considered adequate for the purposes of the study. For the most part, the items on the Observation Record involve the observer in describing the nature of activities, materials and organizational patterns in classrooms - similar to a counting procedure. Some items, admittedly, focus on interaction patterns. However these items are concerned with describing and not evaluating the interaction.

A final limitation was one experienced by most educators who attempt to collect data in the "real world" of the school. This "limitation" refers to the conditions which interfere with collecting all the data desired. In this study, it was necessary to devise and conform to a precise observation schedule. Due to such factors as teacher illness, observer illness, inclement "travel weather" for observers, and situations arising in schools (e.g., special programs), some observations could not proceed as planned on the day scheduled. Further, in most instances, it was not possible to "make up" the observation. As a result, for four of the ten classrooms, only three "full" observations were made (i.e., observation of Conceptual Skills by both observers and observation of remaining program). There was some concern that, especially for Conceptual Skills lessons, scores might be unduly affected by a small number of observations, since a chance occurrence (e.g., a "bad day") could have a larger effect on a teacher's mean score than if five observations had been made. However, inspection of the data ruled this possibility out as far as the present study is concerned, since the scores for these four teachers showed a great deal of consistency.

Although mean scores might have been slightly higher or lower if all observations had been made, the teachers' assignment to the "high" or "low" group would have been the same. It was decided, then, that the small numbers of observations would have little effect, if any, on the results of this study.

The data on classroom observations were more complete. A total of 22 observations were completed in the classrooms of the 5 teachers obtaining the highest mean rating on the Teacher Evaluation form; a total of 22 observations were likewise obtained in classrooms of the 5 teachers obtaining the lowest mean ratings.

Teacher Implementation of Conceptual Skills

Because the items on the Teacher Evaluation form reflect the emphasis of teacher training sessions, the instrument was considered a valid means of assessing teacher performance in implementing the Conceptual Skills Program. Tables 2 and 3 show that when the ratings of the two observers are analyzed by item and by teacher, observer agreement is generally high. Nevertheless Table 2 suggests that although observers were using similar criteria in rating performance on a particular item, agreement was noticeably higher for some items than others. However the degree of observer agreement on item 6, "over all assessment" of what was observed, indicates that observers agreed rather consistently that a teacher's general performance was "good", "adequate", "weak" or "poor" in relation to the purpose and intent of the program.

TABLE 2

OBSERVER AGREEMENT BY ITEM

ITEM	PAIRS OF RATINGS	SAME	%	DIFFERENT	%
1	37	22	59.5	15	40.5
2	37	24	64.9	13	35.1
3	37	27	73.0	10	27.0
4	37	25	67.6	12	32.4
5	37	27	73.0	10	27.0
6	37	28	75.7	9	24.3
TOTAL	222	153	68.9	69	31.1

Table 3 shows that when the ratings of the two observers are analyzed by teacher, observer agreement is higher for some teachers than others. However observer agreement is still generally high for the total sample. The observers' discussion of these results suggested that in their ratings of some teachers on some items, a difference in ratings reflected different observer reaction to such things as "noise level" in the classroom and certain personality traits of the teacher. These observer biases were acknowledged, particularly, in discussing the performance of teachers 303 and 1010.

TABLE 3

OBSERVER AGREEMENT FOR EACH TEACHER

TEACHER	PAIRS OF RATINGS	SAME	%	DIFFERENT	%
101	18	17	94.4	1	5.6
202	18	15	83.3	3	16.7
303	24	10	41.7	14	58.3
404	24	17	70.8	7	29.2
505	18	11	61.1	7	38.9
606	18	14	77.8	4	22.2
707	24	15	62.5	9	37.5
808	24	20	83.3	4	16.7
909	30	21	70.0	9	30.0
1010	24	13	54.1	11	45.8
TOTAL	222	153	68.9	69	31.1

Table 4 shows that when the mean total group ratings are analyzed according to the first 5 items, teachers as a group did not tend to obtain a higher rating on some items than on others. Although there are some differences in the mean ratings of items, these differences are small. Seemingly then there were no specific teaching techniques or procedures which were consistently more difficult or easy for all teachers to implement.

T A B L E 4

MEAN SCORE ON EACH ITEM FOR TOTAL GROUP OF TEACHERS

ITEM	NUMBER OF RATINGS	TOTAL POINTS	$\overline{X}$
1	74	185	2.50
2	74	212	2.86
3	74	202	2.73
4	74	203	2.74
5	74	194	2.62

Table 5 shows that when the mean rating for each teacher on the total of items 1-5 is tabulated, the resulting ranking of teachers supports the hypothesized range of teacher performance in implementing Conceptual Skills considered to be present in the sample. With the exception of one set of tied ranks teacher ratings fall at a different point along the range from generally poor (1.35) to generally good (3.83) program implementation.

Table 6 shows that when a teacher's mean rating on the total of items 1-5 (average of all ratings on all 5 teaching behaviors) is compared with teacher's mean rating on item 6 (average of all ratings of general performance) the mean for item 6 is somewhat higher than the mean for items 1-5 for 6 teachers; somewhat lower for 4 teachers. However inspection of the two sets of means suggests that observers' general assessment of a teacher's performance (Mean, item 6) reflected the assessment derived from averaging the ratings obtained for particular items (Mean, items 1-5). Although each set of means produces a different ranking, what is involved is a change in relative position for some teachers in the group obtaining the 5 highest ratings and some teachers in the group obtaining the 5 lowest ratings. Both sets of rankings permit identification of the same 5 teachers (101, 202, 404, 606, 808) as the "high" group and the same 5 teachers (303, 505, 707, 909, 1010) as the "low" group.

TABLE 5

MEAN SCORE FOR EACH TEACHER ON EACH ITEM
AND TEACHER RANK ON TOTAL 1-5

TEACHER	ITEM 1	2	3	4	5	1-5	RANK
101	3.00	3.33	3.33	3.17	3.33	3.23	2
202	3.50	4.00	3.83	4.00	4.00	3.87	1
303	2.38	2.75	2.50	2.75	2.63	2.60	7
404	2.88	3.63	3.50	2.88	3.13	3.20	3.5
505	2.17	1.83	2.00	1.83	1.83	1.93	9
606	2.50	3.33	3.33	3.33	3.17	3.13	5
707	1.25	1.38	1.50	1.50	1.13	1.35	10
808	2.75	3.50	3.38	3.38	3.00	3.20	3.5
909	2.00	2.40	1.70	2.10	2.10	2.06	8
1010	3.00	2.88	2.88	3.00	2.50	2.85	6

TABLE 6

COMPARISON OF TEACHER RANKINGS ON MEAN SCORE FOR TOTAL OF ITEMS 1-5 AND MEAN SCORE FOR ITEM 6

TEACHER	ITEM 1-5 $\bar{X}$	RANK	TEACHER	ITEM 6 $\bar{X}$	RANK
101	3.23	2	101	3.33	3.5
202	3.87	1	202	3.83	1
303	2.60	7	303	3.00	6
404	3.20	3.5	404	3.13	5
505	1.93	9	505	2.17	8
606	3.13	5	606	3.33	3.5
707	1.35	10	707	1.38	10
808	3.20	3.5	808	3.50	2
909	2.06	8	909	2.00	9
1010	2.85	6	1010	2.63	7

Classroom Observation Record

It is important to note that what is reported concerning the programs of the two teacher groups ("high" and "low") reflects the predominating trends revealed in inspecting the Observation Records for each group. Further the labels "high" and "low" refer respectively and exclusively to high and low ratings of teacher success in implementation of the Conceptual Skills Program.

Programs in the classrooms of the "high" teacher group suggested a generally systematic approach to program planning and implementation. Daily programs appeared to be planned according to a format which gave attention to ordering and balancing activities in a way which established daily routines but, at the same time, permitted variety within each day as well as variety from day to day. Whereas the time of a given activity (e.g., activity period) varied little from day to day, variety was provided within the activities and materials available to children. Another dimension of a systematic approach in these classrooms was revealed in teacher practices of (1) introducing, explaining, discussing "new" materials or equipment, (2) making sure children understood what activities were available to them and/or suggesting activities that could be pursued (3) introducing, reinforcing or reviewing information, concepts, skills through direct instruction which included checking children's understanding.

With respect to program priorities, the programs of teachers in the "high" group suggested concern for (1) helping children to develop "work habits" or skills which contribute to more independent school behaviors (2) helping children to develop language and number concepts and skills as an aspect of developing "readiness" and (3) providing children with a variety of experiences and materials in the areas of music, art and literature. In contrast to programs of teachers in the "low" group, more "readiness" materials and activities were available and more attention was given to instructing children in the use and handling of materials.

In these classrooms these priorities were reflected in the kind as well as the amount of materials and activities available, and in the teacher's approach to working with children. In these respects, it is difficult to separate the suggested priorities from the expectations which seemed to be held for children.

The design and implementation of programs of teachers in the "high" group suggested the point of view that with teacher guidance and support (1) children are capable of assuming some responsibility for their behavior (2) children are capable of learning a number of concepts and skills as preparations for later encounters with the "academic" program of the school (3) children are capable of "choosing" and/or establishing some of their own purposes for learning. Of particular note is the observation that children in these classrooms tended to assume responsibility for such things as getting their own materials, "cleaning up" or putting things away, suggesting or "deciding" what activity could be pursued with what material, choosing "another" activity when one task or activity was completed, responding quickly and easily to teacher requests to join with the total group and getting themselves "ready" to go to the library, the gymnasium or "ready" to go home. For example, in contrast to children in classrooms of the "low" teacher group, children in classrooms of the "high" teacher group required much less time and much less teacher assistance in "dressing" themselves to go out doors.

The programs of teachers in the "high" group might be described in the idiom of the day as being somewhat "structured." The selection of activities and materials available, the ways in which teachers worked with children and the behaviors and responses of children did suggest that programs were planned and implemented according to particular objectives and expectations.

In comparison with programs of teachers in the "high" group, programs of teachers in the "low" group seemed to reflect a less systematic approach to planning and implementation. With respect to ordering or scheduling activities, decisions appeared to be more spontaneous than planned. In contrast to programs of the "high" group, no particular program format was suggested. An activity period

might be provided on one day but not on another. This period might constitute the first activity of the day or be limited to the last 10 minutes of the day. As a rule, daily programs did not provide either the balance, number or variety of activities observed in classrooms of teachers in the "high" group. In terms of working or interacting with children, teachers in the "low" group employed less direct guidance or instruction than teachers in the "high" group. It should be noted, however, that in one of these classrooms, there was a considerable amount of teacher "telling." However the nature of the telling (e.g., telling children what to do, reprimanding children) would not be properly labelled "instruction."

With respect to program priorities, there appeared to be some differences among the teachers in the "low" group. For example, one teacher's chief priority seemed to be permitting children to exercise as much freedom as possible in deciding what to do and when to do it. Child "choice" was generally pursued with a very minimum of teacher guidance or involvement. A second teacher tended to make all choices for all children all of the time. In this latter instance the priority seemed to be getting children to follow directions. In the classrooms of the other teachers in the "low" group it was difficult to detect, from what was observed, what constituted program priorities. The materials, activities and approaches to working with children in these classrooms did not reveal a program pattern or plan which might suggest priorities. There was, however, a consistent feature in all of these classrooms which related to expectations held for children.

Whereas teachers in the "low" group engaged in less teaching or instruction than teachers in the "high" group, they devoted much more time to "waiting on" children and controlling "unacceptable" behavior. Evidently these teachers did not expect children to be responsible for getting their own materials, "cleaning up" after themselves, getting dressed to go outside and the like. These teachers

devoted much time to "servicing" children in these respects. Likewise children in these classrooms seemed to assume little responsibility for other aspects of their behavior. From teacher remarks and reprimands there was some indication that certain "rules" and expectations for behavior has been outlined, but were generally ignored. Of particular note were the numbers of children who either "didn't know" or couldn't decide what they could "do" when permitted a free choice of activity. Such children spent considerable time wandering around the room.

The similarity of program among teachers in the "high" group, did not exist in the same degree among the programs of the teachers in the "low" group. In terms of "structure" both extremes were observed in classrooms of teachers in the "low" group. However in comparison to programs of teachers in the "high" group, the programs of all teachers in the "low" group appeared to have fewer and less clearly defined objectives.

SUMMARY

The evaluation of the Conceptual Skills Program included attention to evaluating student achievement, student attitude, parent attitude, approaches to preparing teacher users of the program, and attempts at program implementation. As a result of this fairly comprehensive evaluation of the program, there is evidence which (1) shows that the program is effective in terms of its objectives, in promoting student achievement (2) indicates student, parent and teacher reaction to the program and (3) suggests what is required for implementing the program successfully. All of these findings are considered by the developers as providing valuable insights relative to the process of producing change in schools.

REFERENCE

Edwards, A.L. Techniques of Attitude Scale Construction. New York: Appleton-Century - Crofts, 1957.

Eisner, E.W. "Curriculum Development: Sources For A Foundation For the Field of Curriculum" in Curriculum Theory Network 5, 1970.

Joyce, B.R. Alternative Models of Elementary Education. New York: Blaisdell Publishing Co., 1969.

Leithwood, K., and Russell, H.R. "A Practical Design for Program Evaluation" in Orbit 11, 3 (1), 1972 pp. 21-22.

APPENDIX A

RATING GUIDE FOR CONCEPTUAL SKILLS OBSERVATIONS

Each observer will give a rating to each of the first five categories. These include:

The teacher involves the entire group.
The teacher supports her children as needed.
The teacher gives her children responsibility.
The teacher enables the children to work at a success level.
The teacher remediates effectively.

Rating of these categories will be according to the following scale:

1. poor, clearly inadequate in handling the items specified under the given category;
2. weak, needs improvement in handling the items specified under the given category;
3. adequate, generally meets basic requirements in handling items specified under given category;
4. good, consistently meets basic requirements in handling items specified under given category.

The rating for the sixth category - The teacher understands the intent of each task and clearly works to achieve its purpose - will indicate the observer's over all assessment of the extent to which the teacher's performance reflects an understanding of the purposes of the program. Rating will be according to the following scale:

1. poor and seemingly distorted understanding of what the program is attempting to achieve with respect to pupil understanding and behavior; seemingly no

attempt to evaluate what is happening during lesson;

2. weak and seemingly only partial understanding of what the program is attempting to achieve with respect to pupil understanding and behavior; few attempts to evaluate what is happening during lesson;

3. adequate understanding of what the program is attempting to achieve with respect to pupil understanding and behavior but needs to improve in evaluating what is happening during lesson;

4. good understanding of what the program is attempting to achieve with respect to pupil understanding and behavior; fairly consistent in evaluating what is happening during lesson.

APPENDIX B

Observation Record

Observer:

Teacher:

Date:

a.m. session: X 9:00-11:15

p.m. session: _____

1. Schedule of activities observed

Opening exercise and Calendar
Conceptual Skills
Activity period
Music (Songs)
Recess
Reading Readiness Activity
Music-Rhythms
Story

2. Proportion of total group, small group, individual activity

Activity period provided opportunities for small group and individual activity. Three small groups of 2-4 children did work and play together during this period. The remaining children worked independently. All other activities were total group oriented.

3. Nature of tasks, materials etc. for individual and/or small group activity (e.g., what is available for children to choose and/or what does teacher outline for individual or small group activity)

The following were available during Activity period:

Art materials including buttons, yarn, cotton, sticks, cardboard boxes, paper, paste.

Paints

Workbench and Housekeeping Corner

Blocks and Block games - e.g., sequence games.

Readiness Worksheets which included many different number oriented sheets - e.g., Drawing a specified number of figures, Identifying given sets of objects, matching number sets etc., a phonics oriented sheet for identifying objects beginning like "dog", classification oriented sheets - e.g., identifying pictures which depict Fall or Spring, identifying objects that belong to mother, father, child etc.

Printing center materials which included models for printing numerals, words and short phrases to be printed - words and phrases related to Spring theme.

4. Selection and performance of individual and/or small group tasks (e.g., is choice child selected or teacher selected, proportion of individual to group activity)

The teacher introduced the Activity period by explaining, to the total group, two "new" readiness worksheets i.e., what the worksheet task required. She then outlined the different activities available and children were allowed to make their own selections. However during the period, the teacher checked to make sure that each child spent a portion of the period at what she terms "hand work" - which seems to include any activity involving materials (e.g., paper, paint, worksheets) as opposed to objects (e.g., games, toys, blocks). Most children engaged in no more than two activities during the period and tended to "stay with" a given activity for a reasonable period of time. Some children spent the entire period with readiness materials; others with art materials. Children at the housekeeping corner, workbench and blocks really worked or played together i.e., planning what they were going to do, interacting with each other etc. This group activity involved approximately 8-10 children, the remaining 10-12 worked independently at tasks.

5. Teacher role during individual/small group activity (e.g., what does the teacher do, how does she interact with children)

The teacher moved from group to group and child to child during the period - asking children what they were doing (or making), suggesting (e.g., Can you think of something else you could put in that picture, something else you see at the lake?) Checking worksheets and remediating errors, printing stories on pictures and the like. When a child sought out the teacher to show her something, she frequently responded with questions or comments which required an "expanded" description or response from the child (e.g., What else can you tell me about this flower you made - What shape is this? Can you tell me something else about this picture? etc.) The teacher also used this period to "check with" individual children relative to incorrect responses given during the Conceptual Skills lesson.

6. Nature of tasks for total group activity (exclusive of Conceptual Skills)

The first Music period included identifying a given song from a few bars played on the piano and then singing the songs. A second activity included singing about and dramatizing a song about riding on a bus. There was some discussion about what comes next (e.g., sequence) on the bus ride.

The Reading Readiness activity involved identifying the beginning sound and/or letter in words presented by the teacher. Model words and picture (e.g., ball, top, moon etc.) were in view of the children and they were encouraged to repeat word the teacher gave and then look for a word on the chart beginning with the same sound or letter.

Since the circus was in town, the teacher presented "circus words" - e.g., tent, balloon etc. The "children" received a point for each correct response. The "teacher" received a point for each incorrect response from the group. The children won the game!!

The second music activity involved listening first to a record to decide what the music "tells us to do" (e.g., skip, walk, march etc.) and then doing it as the record played.

(Over)

6. Cont'd.

The "Story" activity focused on the animals and people in the circus but also focused on number concepts - e.g., 8 clowns, 9 seals etc. The teacher showed each page to the group who identified the people and animals (with some discussion - e.g., what acrobats do, fact that zebras have stripes etc.) and counted the number of people or animals shown.

7. Selection and performance of tasks in total group activities (e.g., who decides what to "do", can child decide not to participate)

The teacher decided what was to be done but children had many opportunities to "participate" i.e., comment, ask questions etc. All children were expected to participate and did.

8. Teacher role in total group activities (e.g., what does the teacher do, how does she interact with children)

The teacher began each total group activity by explaining what was to be done or establishing a purpose for the activity. She gave children many opportunities to respond but kept responses focused on the task at hand. When individual children gave an incorrect or "hazy" response, she "moved in" to remediate the problem - this was especially true in the Reading Readiness activity. She also differentiated her expectations for individual children. For example, children who were not able to identify a beginning letter or sound were encouraged to point to the model picture which "began" like the teacher's "word."